Intelligence 64 (2017) 71–80

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

The Flynn effect for verbal and visuospatial short-term and working memory: A cross-temporal meta-analysis

Peera Wongupparaj[a,b,*], Rangsirat Wongupparaj[c], Veena Kumari[a,d], Robin G. Morris[a,e]

[a] Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK
[b] College of Research Methodology and Cognitive Science, Burapha University, Thailand
[c] Faculty of Education, Chulalongkorn University, Thailand
[d] NIHR Biomedical Research Center for Mental Health, South London and Maudsley NHS Foundation Trust, London, UK
[e] Department of Clinical Neuropsychology, Clinical Neurosciences, King's College Hospital NHS Foundation Trust, London, UK

ARTICLE INFO

Keywords:
Flynn effect
Short-term memory
Working memory
Forward and backward digit span
Forward and backward Corsi block span
Cross-temporal meta-analysis

ABSTRACT

The Flynn effect has been investigated extensively for IQ, but few attempts have been made to study it in relation to working memory (WM). Based on the findings from a cross-temporal meta-analysis using 1754 independent samples ($n = 139{,}677$), the Flynn effect was observed across a 43-year period, with changes here expressed in terms of correlations (coefficients) between year of publication and mean memory test scores. Specifically, the Flynn effect was found for forward digit span ($r = 0.12$, $p < 0.01$) and forward Corsi block span ($r = 0.10$, $p < 0.01$). Moreover, an anti-Flynn effect was found for backward digit span ($r = -0.06$, $p < 0.01$) and for backward Corsi block span ($r = -0.17$, $p < 0.01$). Overall, the results support co-occurrence theories that predict simultaneous secular gains in specialized abilities and declines in *g*. The causes of the differential trajectories are further discussed.

1. Introduction

The phenomenon called the Flynn Effect (FE) is the increase in cognitive test performance in a population over time, as originally observed by Runquist (1936), who found a systematic increase in scores on the Minnesota College Aptitude test between 1929 and 1933. The FE term was coined by Herrnstein and Murray (1994, p. 307), after James Flynn who demonstrated specifically increases in the Intelligence Quotient (IQ) (Flynn, 1984, 1987, 1990, 2012), also confirmed by a range of researchers in this field (e.g. Pietschnig & Voracek, 2015; Pietschnig, Voracek, & Formann, 2010; Trahan, Stuebing, Hiscock, & Fletcher, 2014).

Apart from the FE for IQ, secular gains for working memory (WM), which can be defined as the immediate and transient retention of information for temporary storage and processing, might be also expected given the strong link between this function and IQ (e.g. Ackerman, Beier, & Boyle, 2005; Cohen & Sandberg, 1977; Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Colom, Abad, Rebollo, & Chun Shih, 2005; Colom, Flores-Mendoza, Quiroga, & Privado, 2005). Indeed, tests of WM such as forward and backward digit span have often been incorporated in IQ test batteries, including the various versions of the Wechsler Adult Intelligence Scale (WAIS) (Wechsler, 1981, 1987, 2008). Here, a common procedure is the forward digit span test, in which strings of digits are presented for immediate recall, the lengths of strings increasing until the person consistently makes errors, with the backward digit span variant involving reverse order recall of the digit strings (Woods et al., 2011). As might be predicted, a WM FE has been reported using digit span tasks, albeit with inconsistent findings across studies. For example, Wicherts et al. (2004) compared digit span scores between two Dutch standardization cohorts, one from 1967/1968 ($n = 2100$) and the other from 1998/1999 ($n = 77$). The results revealed the smallest gain was on a digit span subscale (effect size = 0.50; S.D. units) in comparison to the highest gain on a similarity subscale (effect size = 1.48; S.D. units). However, Kanaya and Ceci (2011) also investigated a longitudinal FE on digit span task among 672 school children who were assessed in 1974 and again in 2002, finding no difference in scores. Recently, Gignac (2015) analyzed forward and backward digit span data from 10 normative studies conducted between 1923 and 2008 and showed that there were no significant trends in any of them (respectively, $r_{\text{forward}} = 0.45, p = 0.27$ and $r_{\text{backward}} = -0.57, p = 0.12$; $k = 9$). However, shortly afterwards, Woodley of Menie and Fernandes (2015) re-analyzed Gignac's (2015) data with the same 10 normative samples but using a weighted regression by sample size method. This result produced modest to large associations ($r_{\text{forward}} = 0.58$ and $r_{\text{backward}} = -0.55$; $n = 6841$), their analysis was the first study to demonstrate robust and opposing

---

* Corresponding author at: College of Research Methodology and Cognitive Science, Burapha University, Chonburi Province 20131, Thailand.
E-mail addresses: Peera.1.wongupparaj@kcl.ac.uk (P. Wongupparaj), Veena.kumari@kcl.ac.uk (V. Kumari), Robin.morris@kcl.ac.uk (R.G. Morris).

http://dx.doi.org/10.1016/j.intell.2017.07.006
Received 11 March 2016; Received in revised form 10 July 2017; Accepted 17 July 2017

temporal trends for these indicators. Accordingly, Woodley of Menie and Fernandes (2015)'s statistical technique was thought superior to Gignac's method in terms of statistical power to detect the effect. However, both studies have limitations due to using only a relatively small number of studies (i.e. k = 10). Moreover, although verbal WM as indexed by backward digit span tasks has been widely used for a long time in both clinical and non-clinical settings, there are still very few studies that have looked for FE using verbal WM. Thus, further research on this topic with a more substantial sample size is needed.

In parallel, studies have shown a strong relationship between visuospatial WM and cognitive abilities (e.g. Ashkenazi, Rosenberg-Lee, Metcalfe, Swigart, & Menon, 2013; Colom et al., 2005; Colom et al., 2005; Kyttälä & Lehto, 2008; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). Despite the fact that non-verbal abilities tend to show a larger FE (Pietschnig & Voracek, 2015; Rönnlund, Carlstedt, Blomstedt, Nilsson, & Weinehall, 2013), to our knowledge, the FE for visuospatial WM has received very little attention and no study has yet been reported. The Corsi block task has long been used as a key measure for visuospatial WM since it was created by Corsi in 1972 (Corsi, 1972; Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000). This task consists of an array of nine cubical blocks in which the participant has to remember sequences of blocks "tapped" out by an experimenter. As with digit span, the task starts with short sequences and terminates at higher levels when performance is consistently unsuccessful. The participant is required to repeat the given sequence either identically or in the reversal order (Brunetti, Del Gatto, & Delogu, 2014a; Kessels et al., 2000). Since it has been widely used for a long period of time, this task is a rich resource for conducting meta-analyses to address the gap in our understanding concerning a visuospatial WM FE.

Accordingly, the objectives of the present study were to investigate the FE and the anti-FE for both verbal and visuospatial WM using a Cross-Temporal Meta-Analysis (CTMA). This technique provided a possibly beneficial way to observe the change tendency of a set of specific variables over a period of time and to test possible mediators underlying that change (Donnelly & Twenge, 2017; Zhang, Tan, Wu, Han, & Wang, 2016). In this study, the term short-term memory (STM) was used to indicate working memory functioning in which only temporary storage was required, as measured using forward digit span and forward Corsi block, respectively, for verbal and visuospatial STM. This is contrasted with WM, here defined as involving both storage and processing of information, in this study backward digit span and backward Corsi block span used respectively as measures of verbal and visuospatial WM.

2. Methods

2.1. Literature search

The target articles were identified through 15 scientific Databases from 1910 to 2016; i) Taylor & Francis Online, ii) ScienceDirect, iii) SpringerLink, iv) PsycARTICLES & PsychINFO, v) Wiley Online Library, vi) Sage Journals, vii) Oxford Journals, viii) Cambridge Journals, ix) ProQuest dissertations and theses, x) Karger Online Journals, xi) BMJ, xii) JSTOR, xiii) Frontiers in Psychology, xiv) BioMed Central, and xv) HighWire Journal. The following search terms were used individually and in combination using the Boolean OR function to maximize search sensitivity; "digit span*", "Wechsler's digit span", "forward digit span", "backward digit span", "Corsi block*", "forward Corsi block", and "backward Corsi block" (* is a truncation symbol to represent multiple endings and spellings). No limits were set. The search also identified any data published in journals as well as that found in theses or dissertations (grey literature or unpublished material). Results from all e-databases were combined and duplicates removed subsequently. Despite the fact that we included doctoral dissertations in the current dataset, others sources of grey literature were not systematically searched (e.g. unpublished research reports, submitted or in-press articles). This may have resulted in some related data not being included in this CTMA.

2.2. Inclusion and exclusion criteria

The meta-analysis included data from studies specifically using the Wechsler-Bellevue scales and its variants for forward/backward digit span and those using the standard versions of the forward/backward Corsi block tasks (see Corsi, 1972; Kessels et al., 2000). Included in the final data-set were data only from studies that reported the mean and/or standard deviation of the raw score for 'longest forward or backward spans' (LFS or LBS) recalled without error. Thus, studies were excluded if they reported only total digit span or Corsi block span scores, the scores for number of items (trials) correct or standardized scores as well as age-scaled scores. Also, studies had to use the same design protocols as the original version of the test (Corsi, 1972) in which the number of blocks increased from a minimum of 2 to a maximum of 9 blocks and the participant was given two attempts at each pattern length. In addition, if studies adopted a test–retest method, only the mean and/or standard deviation scores for pre-test or baseline were recorded and if several articles investigated the same sample or used the same dataset, the statistical parameters (means, standard deviations, standard errors) were treated as a single data point. Studies were excluded if they investigated clinical research participants (with mental or physical disabilities) with the exception of when they reported data for control groups, as such data are appropriate for use in this study. Also, review articles, systematic reviews, meta-analyses, research protocols and case report studies were excluded due to the fact that no targeted statistical parameters are reported in these studies.

In short, there were four main memory tests considered, namely Forward Digit Span (FDS), Backward Digit Span (BDS), Forward Corsi Block (FCB), and Backward Corsi Block (BCB); these were designated in the meta-analysis as representing respectively verbal short-term memory (FDS), visuospatial short-term memory (FCB), verbal working memory (BDS), and visuospatial working memory (BCB).

2.3. Moderator variables

Age was used as a moderator variable, because of the possibility that the mean age for each study would account for variations in the relationship between mean scores for memory tests and years of publication, given evidence for aging affecting STM and WM capacities (Bopp & Verhaeghen, 2005; Hester, Kinsella, & Ong, 2004; Myerson, Emery, White, & Hale, 2003; Pelosi & Blumhardt, 1999; Wingfield, Stine, Lahar, & Aberdeen, 1988). Similarly, gender was also used as a moderator, with previous studies showing gender differences on memory tests, with male participants tending to produce higher scores on spatial tasks but female participants tending to have higher scores on verbal tasks (Andreano & Cahill, 2009; Hill, Laird, & Robinson, 2014; Lowe, Mayfield, & Reynolds, 2003). Thus, the relative strengths for male and female participants on memory tasks may confound the interpretation of the FE on STM and WM tasks.

In addition, prior research has suggested possible negative effect of malnutrition, poorer health, and poverty on cognitive abilities of people in developing countries (Grantham-McGregor et al., 2007), especially for memory development (Khor & Misra, 2012). Furthermore, many studies have shown lower fertility rate among women in developed countries than in developing countries (Bureau, 2016; Nargund, 2006). The FE and anti-FE were also found as partly being caused by fertility patterns in many countries (Kong et al., 2017; Menie, Fernandes, José Figueredo, & Meisenberg, 2015; Sundet, Borren, & Tambs, 2008). Hence the study investigated whether or not this aspect affects memory scores of people between the two types of countries over time. Here, type of country was divided into "developing" versus "developed" according to the terminology recently adopted by the International Statistical Institute (ISI) (2016). A further consideration was the test platform of the

Table 1
Sample characteristics.

Variables	FDS (ks = 742) (ns = 48,955)		BDS (ks = 594) (ns = 70,424)		FCB (ks = 307) (ns = 16,514)		BCB (ks = 111) (ns = 3784)	
	Mean (SD)	Min – Max	Mean (SD)	Min – Max	Mean (SD)	Min – Max	Mean (SD)	Min – Max
Years of publication[a]	2003.15 (15)	1973–2016	2005.33 (15)	1975–2016	2008.22 (9)	1982–2016	2009.71 (6)	1989–2016
Age of participant[b]	36.76 (230.90)	3.65–92.40	41.27 (287.11)	4–92.50	29.31 (162.48)	3.04–86.65	43.06 (149.39)	4–86.65
Mean score[b]	6.11 (7.22)	0.81–9.10	4.51 (9.88)	1.08–7.20	5.25 (6.35)	2.20–7.80	4.74 (4.81)	2.10–7.09
SD score[b]	1.33 (3.41)	0.10–3.08	1.38 (4.34)	1.08–7.20	1.11 (3.25)	2.20–7.80	1.01 (1.90)	0.17–2.53
Sex[b]								
Male[c]	5.77 (9.37)	3.70–7.80	4.30 (5.67)	2.10–6.90	4.86 (5.90)	3.10–7.00	5.02 (3.40)	2.70–5.70
Female[d]	6.12 (10.76)	3.40–8.74	4.64 (6.86)	2.10–7.13	4.80 (5.92)	3.70–7.00	4.75 (4.90)	2.10–5.93
Sample size	65.98 (191.85)	1–4100	118.56 (352.58)	5–4251	53.79 (80.52)	9–632	34.09 (39.56)	10–362
Test platform								
Non-computerized (paper-and-pencil or wooden cube) (%)	676 (91.10)	–	552 (92.90)	–	229 (74.60)	–	99 (89.20)	–
Computerized (%)	66 (8.90)		42 (7.10)		78 (25.40)		12 (10.80)	
Type of country								
Developed	673 (90.70)	–	529 (89.10)	–	283 (92.20)	–	87 (78.40)	–
Developing	69 (9.30)		65 (10.90)		24 (7.80)		24 (21.60)	

ks = Number of independent samples in which means memory test was available for the analysis, ns = Sum of the sample size for each memory test.

FDS = Forward Digit Span; BDS = Backward Digit Span; FCB = Forward Corsi Block; FCB = Backward Corsi Block.

[a] Median and interquartile range (IQR).

[b] Sample size weights.

[c] Weighted mean with 95% Confidence Interval for male; FDS = 5.77 (5.50–6.04); BDS = 4.30 (4.11–4.50); FCB = 4.86 (4.63–5.10); BCB = 5.02 (4.53–5.51).

[d] Weighted mean with 95% Confidence Interval for female; FDS = 6.12 (5.82–6.43); BDS = 4.64 (4.37–4.91); FCB = 4.80 (4.59–5.01); BCB = 4.75 (4.20–5.30).

memory tests used and this was included as a moderator in the current analysis. The test platform could be categorized into computerized and non-computerized versions, as shown in Table 1.

2.4. Statistical analyses

The study used the modified meta-analysis technique, CTMA, to analyze the extent to which the mean scores for memory tests changed over decades based on temporal correlation coefficients, utilizing year of study publication and weighted based on sample size (Twenge, 2000). To estimate the magnitude of change among the mean scores over time, a moderated and weighted least squares multiple regression was applied using Statistica 13 and Statgraphics. Both standardized and unstandardized regression weights were reported. The mean age of participant, percentage of men in each study, test platform (computerized vs. non-computerized), and type of country (developed vs. developing) were inputted into the model as moderators. To indicate the magnitude of change in STM and WM scores over time, a weighted linear regression was calculated for each memory test. The year of publication was used as the predictor with the sample mean FDS, BDS, FCB or BCB scores as the outcome measures (Pietschnig et al., 2010).

In the current analysis, the number of individuals (n) was used to estimate the model degrees of freedom and significances. Essentially, the CTMA method is used to analyze trend scores over time (e.g. Twenge, 2000; Twenge & Campbell, 2001; Twenge & Campbell, 2008; Wongupparaj, Kumari, & Morris, 2015). In contrast to other types of meta-analysis, the primary outcome of interest is not the effect size. The temporal relationship between mean score and year of publication of these scores were collected from target studies and analyzed. In addition, this technique was used to analyze sample mean data, also weighting the analysis by sample size from each study, so taking into account the fact that bigger sample sizes should be weighted more highly and hence accurately (see Lipsey & Wilson, 2001).

In a conventional meta-analysis of effect sizes, the validity of the research findings may be affected by publication bias, namely the file drawer effect (Rothstein, Sutton, & Borenstein, 2006). However, a factor mitigating this effect is that the mean scores from each published study was collected and analyzed, rather than the effect sizes from the primary measures of the study. This related to the target mean score for the digit span or Corsi block measures rarely being associated with the central hypotheses considered by the journal papers, tending to be reported on an incidental basis as descriptive or Supplementary data (Booth, Sharma, & Leader, 2016; Cooper, 1998; Pietschnig et al., 2010; Schmidt & Hunter, 2015).

Initially, outlier and influential case diagnostics were performed using Cook's distances, DFBETAS, and DFFITS (Viechtbauer & Cheung, 2010). First, according to Skrondal and Rabe-Hesketh (2004), for the value of Cook's distances to be “large” they should be four times the number of parameters in the model divided by the number of independent samples (i.e. > 0.01, > 0.01, > 0.03, and > 0.07 for FDS, BDS, FCB, and BCB, respectively). Second, data were considered outliers if the value of the absolute DEFETAS exceeding $2/\sqrt{n}$, where n is the number of independent sample (i.e. > 0.07, > 0.08, > 0.11, and > 0.19 for FDS, BDS, FCB, and BCB, respectively). Third, outliers were also defined as when the value of the absolute DFFITS exceeded $2\sqrt{p/n}$, where p is the number of parameter in the model and n is the number of independent sample (i.e. > 0.10, > 0.12, > 0.16, and > 0.27 for FDS, BDS, FCB, and BCB, respectively). Applying these principles, outliers being detected from 30 FDS, 12 BDS, 10 FCB, and 6 BCB studies and data from these studies were then deleted from the final dataset. Following removal of the outliers, some changes in the observed relationship between the mean memory test scores and years of publication were found ($r_{\text{FDS with vs without outlier}} = 0.12^{**}$ vs 0.12^{**}, $r_{\text{BDS with vs without outlier}} = -0.10^{**}$ vs -0.06^{**}, $r_{\text{FCB with vs without outlier}} = 0.13^{**}$ vs 0.10^{**}, and $r_{\text{BCB with vs without outlier}} = 0.01^{\text{ns}}$ vs -0.17^{**}). However, these numerical changes concerning r_{FDS}, r_{BDS}, r_{FCB}, and r_{BCB} showed no significant differences ($z_{\text{FDS}} \leq 0.01$, $p = 1.00$; $z_{\text{BDS}} = -0.70$, $p = 0.48$; $z_{\text{FCB}} = 0.38$, $p = 0.70$; and $z_{\text{BCB}} = 1.38$, $p = 0.17$) and the effect sizes were small ($d_{\text{FDS}} \leq 0.01$, $d_{\text{BDS}} = 0.05$, $d_{\text{FCB}} = 0.04$, and $d_{\text{BCB}} = 0.25$) (see Supplementary information: S1).

In addition, the findings from the current dataset as to publication status (i.e. weighted means for memory tests from published vs. unpublished sources) did not show any significant differences and the magnitudes of difference as indicated by effect sizes (i.e. partial η^2) were small (Cohen, Miles, & Shevlin, 2001) for FDS ($p = 0.15$, partial $\eta^2 < 0.01$), BDS ($p = 0.14$, partial $\eta^2 < 0.01$), and BCB ($p = 0.49$, partial $\eta^2 < 0.01$), respectively. Nevertheless, the weighted mean score for FCB indicated a significant difference between published and unpublished sources but the effect size was also small ($p < 0.05$, partial $\eta^2 = 0.04$).

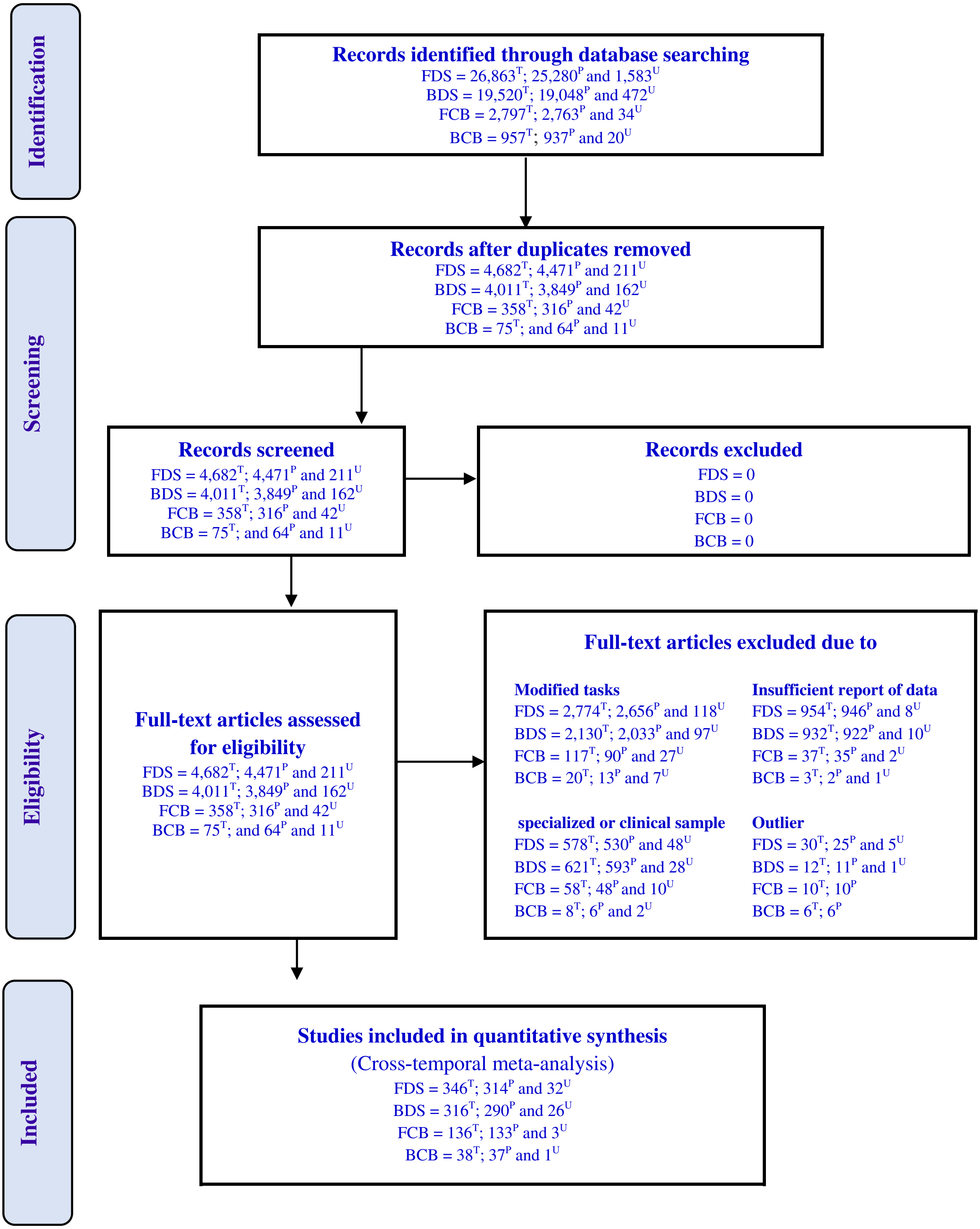

Fig. 1. The PRISMA diagram with literature search results.

Note. FDS = Forward Digit Span, BDS = Backward Digit Span, FCB = Forward Corsi Block, and BCB = Backward Corsi Block. T = Total, P = Published, and U = Unpublished studies.

2.5. *Final sample*

The literature review processes according to PRISMA guidelines are shown in Fig. 1. The 15-electronic database search yielded 26,863 studies for FDS (94.11% published source), 19,520 studies for BDS (97.58% published source), 2797 studies for FCB (98.78% published source), and 957 studies for BCB (97.91% published source). Of these, duplicated and unrelated studies were removed after title, keyword, and abstract review. Only, 4682 (95.49% published source), 4011 (95.96% published source), 358 (88.27% published source), and 75 studies (85.33% published source) for FDS, BDS, FCB, and BCB remained for the next step. Subsequently, 346 (90.75% published source), 316 (91.77% published source), 136 (97.79% published source), and 38 (97.37% published source) studies for FDS, BDS, FCB, and BCB respectively fulfilled the inclusion and exclusion criteria, resulting in 742, 594, 307, and 111 independent samples for FDS, BDS, FCB, and BCB) and 139,677 participants overall (48,955, 70,424, 16,514, and 3784 participants for FDS, BDS, FCB, BCB). References for studies included in this CTMA can be found in the supporting information and the dataset file is openly available at https://osf.io/4zesc/.

3. Results

Descriptive statistics concerning years of publication, age of participant, sex, sample size, test platform, and type of country are given in Table 1 for each memory test. The median years of publication were as follows: for FDS, 2003.15 (IQR = 15; span 43 years; 1973–2016), for BDS 2005.33 (IQR = 15; span 41 years, 1975–2016), for FCB, 2008.22 (IQR = 9; span 34 years, 1982–2016), and for BCB 2009.71 (IQR = 6; span 27 years, 1989–2016). The weighted mean ages of the participants were; FDS: 36.76 (SD = 230.90; span 3.65–92.40 years), BDS: 41.27 (SD = 287.11; span 4.00–92.50), FCB: 29.31 (SD = 162.48; span 3.04–86.65); and BCB: 43.06 (SD = 149.39; span 4.00–86.65). The weighted mean scores were; FDS: 6.11 (SD = 7.22); BDS: 4.51 (SD = 9.88); FCB: 5.25 (SD = 6.35); and BCB: 4.74 (SD = 4.81). In addition, the weighted means of SD were; FDS: 1.33 (SD = 3.41); BDS: 1.38 (SD = 4.34); FCB: 1.11 (SD = 3.25); and BCB: 1.01 (SD = 1.90).

Moreover, the weighted mean scores for the four memory tests additionally demonstrated that female participants had higher mean scores than male participants for FDS and BDS (6.12 vs. 5.77 and 4.64 vs. 4.30), but male participants had higher mean scores than female participants for FCB and BCB (4.86 vs. 4.80 and 5.02 vs. 4.75). The average sample sizes for FDS, BDS, FCB and BCB respectively were 65.98 (SD = 191.85), 118.56 (SD = 352.58), 53.79 (SD = 80.52), and 34.09 (SD = 39.56). The most frequent types of test platform for all STM and WM tests were non-computerized versions (for FDS, BDS, FCB and BCB respectively, 91.10%, 92.90%, 74.60%, and 89.20%) and the majority of data for all memory tests were collected from developed countries (respectively 90.70%, 89.10%, 92.20%, and 78.40%).

To test the hypotheses that STM and WM increase or decrease overtime, the correlation coefficients from four weighted simple regression analysis were estimated (reported in Tables 2A and 2B and Fig. 2). It was found that the mean scores for the FDS and FCB measures correlated positively with year of publication (respectively: $r = 0.12$, $p < 0.01$; $r = 0.10$, $p < 0.01$; both for sample size weight). A negative correlation was found between the mean score for BDS and BCS and year of publication (respectively; $r = -0.06$ for sample size weight, $p < 0.01$; $r = -0.17$ for sample size weight, $p < 0.01$; both for sample size weight).

The analysis tested the moderating effects of the relationship between the mean scores for all memory tests and year of publication, the main finding (see Tables 2A and 2B) being that the moderators were significant predictors for all memory tests. Specifically, the significant correlations between age of participant and FDS ($\beta_{FDS} = 0.14$, $p < 0.01$), BDS ($\beta_{BDS} = 0.17$, $p < 0.01$), and FCB ($\beta_{FCB} = 0.06$, $p < 0.01$) mean scores indicated higher test performances for older participants. Nonetheless, a negative and significant relationship was found between age of participant and the BCB mean score ($\beta_{BCB} = -0.45, p < 0.01$), this indicating higher test performance for younger participants. Further, a positive and significant correlation was observed between male percentage and the BCB means score ($\beta_{BCB} = 0.10$, $p < 0.01$), showing higher test performances in male participants on this memory test. However, negative directions were found for the relationships between male percentage and FDS ($\beta_{FDS} = -0.03, p < 0.01$), BDS ($\beta_{BDS} = -0.01, p < 0.01$), and FCB ($\beta_{FCB} = -0.06, p < 0.01$) mean scores, indicating better performance for female participants on these three memory tests.

Significant negative relationships between test platform and FDS ($\beta_{FDS} = -0.02$, $p < 0.01$), BDS ($\beta_{BDS} = -0.03$, $p < 0.01$), FCB ($\beta_{FCB} = -0.08, p < 0.01$), and BCB ($\beta_{BCB} = -0.14, p < 0.01$) mean scores suggested higher test score of the computerized platform in comparison to the non-computerized platform on all four memory tests. Also, it is clear from the current analysis that positive and significant correlations were found between type of country and FDS ($\beta_{FDS} = 0.16$, $p < 0.01$), BDS ($\beta_{BDS} = 0.26$, $p < 0.01$), FCB ($\beta_{FCB} = 0.07$, $p < 0.01$), and BCB ($\beta_{BCB} = 0.30, p < 0.01$) mean scores, indicating better test scores for the participants from developed countries than those from developing countries on the four memory tests. In addition, according to the multiple weighted regression, the relationships between years of publication and mean scores were still statistically significant for all memory tests ($\beta_{FDS} = 0.14, p < 0.01, \beta_{BDS} < -0.02$, $p < 0.01, \beta_{FCB} = 0.08, p < 0.01$, and $\beta_{BCB} = -0.03, p < 0.01$) after the moderating effects of participant age, male percentage, type of platform, and type of country were taken into account.

4. Discussion

The aim of the study was to determine whether the mean scores on

Table 2A

Single and multiple weighted least squares regression analyses as weighted by sample size for forward digit span and backward digit span scores and score changes per decade.

Target variables	Ns	b (SE)	β (SE)	*t*	*p*	*r* (R^2)
(1) FDS						
Single regression	48,955					0.12 (0.02)**
Intercept		− 17.43 (0.85)		− 20.40**	< 0.01	
Year of publication		0.01 (< 0.01)	0.12 (< 0.01)	27.56**	< 0.01	
Multiple regression	48,995					0.26 (0.07)**
Intercept		− 21.16 (0.85)		− 24.89**	< 0.01	
Year of publication (Y)		0.01 (< 0.01)	0.14 (< 0.01)	31.36**	< 0.01	
Mean age (A)		< 0.01 (< 0.01)	0.14 (< 0.01)	30.91**	< 0.01	
Male percentage (M)		< − 0.01 (< 0.01)	− 0.03 (< 0.01)	− 6.94**	< 0.01	
Platform (P)		< − 0.05 (< 0.01)	< − 0.02 (< 0.01)	− 3.88**	< 0.01	
Type of country (TC)		< 0.67 (0.02)	0.16 (< 0.01)	36.97**	< 0.01	
(2) BDS						
Single regression	70,424					− 0.06 (< 0.01)**
Intercept		16.11 (0.76)		21.24**	< 0.01	
Year of publication		− 0.01 (< 0.01)	− 0.06 (< 0.01)	− 15.29*	< 0.01	
Multiple regression	70,424					0.32 (0.10)**
Intercept		7.74 (0.73)		10.52	< 0.01	
Year of publication (Y)		< − 0.01 (< 0.01)	− 0.02 (< 0.01)	− 5.67**	< 0.01	
Mean age (A)		< 0.01 (< 0.01)	0.17 (< 0.01)	45.20**	< 0.01	
Male percentage (M)		< − 0.02 (< 0.01)	− 0.01 (< 0.01)	− 4.60**	< 0.01	
Platform (P)		− 0.09 (0.01)	− 0.03 (< 0.01)	− 7.59**	< 0.01	
Type of country (TC)		0.88 (0.01)	0.26 (0.01)	73.20**	< 0.01	

Table 2B
Single and multiple weighted least squares regression analyses as weighted by sample size for forward Corsi block span and backward Corsi block span scores and score changes per decade.

Target variables	Ns	b (SE)	β (95% CI)	t	p	r (R^2)
(1) FCB						
Single regression	16,514					0.10 (0.01)**
Intercept		−13.47 (1.50)		−8.98**	< 0.01	
Year of publication		0.01 (< 0.01)	0.10 (< 0.01)	12.48**	< 0.01	
Multiple regression	16,514					0.18 (0.03)**
Intercept		−11.44 (1.54)		−7.44**	< 0.01	
Year of publication (Y)		0.01 (< 0.01)	0.08 (0.01)	10.75**	< 0.01	
Mean age (A)		< 0.01 (< 0.01)	0.06 (0.01)	7.38**	< 0.01	
Male percentage (M)		< −0.01 (< 0.01)	−0.06 (0.01)	−8.34**	< 0.01	
Platform (P)		−0.15 (0.01)	−0.08 (0.01)	−10.14**	< 0.01	
Type of country (TC)		0.36 (0.04)	0.07 (0.01)	9.64**	< 0.01	
(2) BCB						
Single regression	3784					−0.17 (0.03)**
Intercept		46.54 (3.81)		12.20	< 0.01	
Year of publication		−0.02 (< 0.01)	−0.17 (0.02)	−10.96**	< 0.01	
Multiple regression	3784					0.51 (0.26)**
Intercept		12.45 (3.56)	−0.03 (0.02)	3.50	< 0.01	
Year of publication (Y)		< −0.01 (< 0.01)	−0.45 (0.02)	−2.13**	< 0.01	
Mean age (A)		< −0.01 (< 0.01)	0.10 (0.01)	−28.50**	< 0.01	
Male percentage (M)		< 0.01 (< 0.01)	−0.14 (0.02)	7.04*	< 0.01	
Platform (P)		−0.39 (0.04)	0.30 (0.01)	−9.51**	< 0.01	
Type of country (TC)		0.73 (0.04)	−0.03 (0.02)	20.16**	< 0.01	

Note Ns = number of participants, *$p < 0.05$, and **$p < 0.01$.

tests of verbal and visuospatial STM and WM change over decades, in line with the FE and the anti-FE. Accordingly, the FE for STM and anti-FE for WM were investigated using CTMA on 1754 independent samples. The results revealed the FE for the verbal and visuospatial STM measures (i.e. FDS and FCB) and anti-FE for the verbal and visuospatial WM measure (i.e. BDS and BCB). Both verbal and visuospatial STM measures are possibly sensitive to the FE in which the mean scores for FDS and FCB increase over time ($rs = 0.12$ and 0.10, $ps < 0.01$). In contrast, both verbal and visuospatial WM measures are probably sensitive to the anti-FE in that the mean scores for BDS and BCB decrease over time ($rs = -0.06$ and -0.17, $ps < 0.01$). The gain on the verbal and visuospatial STM tasks (i.e. FDS and FCB) and decline on the verbal and visuospatial WM tasks (i.e. BDS and BCB) are consistent with the co-occurrence model (Woodley of Menie & Fernandes, 2015). Based on this model, it is predicted that less *g*-loaded measures should show a rise, whereas more *g*-loaded measures may show the opposite trend as shown in the studies by Gignac (2015) and Woodley of Menie and Fernandes (2015).

Some studies have also supported the notion that FDS and BDS are differentially related to *g*, namely that BDS is essentially a more attentionally demanding task than FDS, such that BDS is more strongly related to *g* (Ganzach, 2016; Jensen & Figueroa, 1975). In addition, the less *g*-loaded indicator implies a major environmental influence, with the FE gain predominantly driven by environmental causes (te Nijenhuis & van der Flier, 2013), with additional practice and training effects (Bartels, Wegrzyn, Wiedl, Ackermann, & Ehrenreich, 2010; Chase & Ericsson, 1981; Roediger & Karpicke, 2006; St Clair-Thompson & Holmes, 2008). Furthermore, the existence of the FE for STM tasks (i.e. FDS and FCB) possibly adds robustness to the overall FE phenomenon, as it demonstrates in this study that the FE occurs on ratio-scale measures, and does not simply emerge as an artefact of comparing between cohorts measured on interval scale measures of cognitive ability (Woodley of Menie, Peñaherrera, Fernandes, Becker, & Flynn, 2016).

In comparison to the FDS task, the performance score on the more *g*-loaded BDS measure may be potentially related to genetic-driven factors (Ando, Ono, & Wright, 2001; Karlsgodt, Bachman, Winkler, Bearden, & Glahn, 2011; Karlsgodt et al., 2010; Kremen et al., 2007). Galton (1869) was the first to propose the existence of dysgenic effects on cognitive abilities and this has been considered in detail by Lynn (2011), with Woodley of Menie and Fernandes (2015) suggesting that the negative trend for BDS score concerning year of publication could

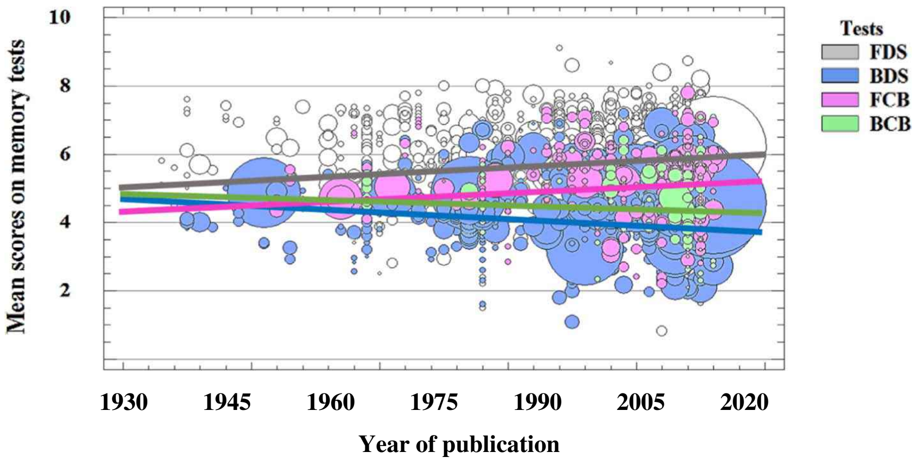

Fig. 2. Weighted simple regression of the relationship between the mean scores of four memory tests (FDS, BDS, FCB, and BCB) and year of publication. The bubble colours represent different types of memory tests, and the sizes provide a visual analogue of the relative sample size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

also be explained by dysgenic fertility on the more *g*-loaded tasks, as found in several studies (Peach, Lyerly, & Reeve, 2014; Woodley of Menie & Dunkel, 2015; Woodley of Menie, Fernandes, Figueredo, & Meisenberg, 2015; Woodley of Menie, Figueredo, Dunkel, & Madison, 2015; Woodley of Menie & Meisenberg, 2013). That is, it has been proposed that more intelligent men and women in developed countries tend to have fewer offspring and also more highly educated woman are more likely to desire childlessness (Kanazawa, 2014; Livingston & Cohn, 2010; Meisenberg, 2010). Speculatively, this might support a steady decline in cognitive ability in some advanced industrial nations (Kanazawa, 2014). Using data from a handful of developed and developing countries, several attempts have been made to estimate the theoretical IQ lost due to the action of dysgenic selection, these countries including the UK and US (for a meta-analysis see Woodley of Menie, 2015), Taiwan (Chen, Chen, Liao, & Chen, 2013), and Kuwait (Abdel-Khalek & Lynn, 2008). Furthermore, anti-FEs on pencil-and-paper tests have also been observed across several different ability measures and in several countries (for reviews, see Dutton & Lynn, 2013, 2015). These IQ declines, like dysgenic effects, tend to be biggest when estimated use more *g* loaded subtests (Woodley & Meisenberg, 2013; Woodley of Menie & Dunkel, 2015).

Moreover, there is also evidence from the meta-analysis that the WM span is significantly more age sensitive than the STM span (Bopp & Verhaeghen, 2005), suggesting that WM is more susceptible to decline than STM with advanced age, confirming previous suggestions (Bowles & Salthouse, 2003; Hale et al., 2011; McNab et al., 2015; Salthouse, 1994; Wingfield et al., 1988). The increasing longevity in developed countries has resulted in a higher proportion of aging population (Ediev, 2011). For instance, mortality at age 75 years or more in the UK accounted for approximately 12% of all deaths of the last century; and it was around 39% in 1951, dramatically rising in the 21st Century to 65% in 2004 (Howse, 2006). In the current dataset, the samples were composed of 29.9% from ‘ $>$ 60 years old’, followed by 21.2% from ‘13–25 years old’, and 19.6% from ‘7–12 years old’, respectively. Based on a lifespan perspective of age-related cognitive decline, it is plausible that WM decline over time may stem in part from the performance scores from this older age group across the more economically developed nations.

In addition, with respect to the negative relationship between visuospatial WM, as indexed by BCB, and year of publication, this negative change also supports the co-occurrence model (Woodley of Menie & Fernandes, 2015). Visuospatial WM has been found to be a strong predictor of mathematical ability (e.g. Ashkenazi et al., 2013; De Smedt et al., 2009; Hubber, Gilmore, & Cragg, 2014; van der Ven, van der Maas, Straatemeier, & Jansen, 2013). The mathematic scores on the International Mathematical Olympiad (IMO) during 1959–2015 indicate a significant and negative trend between the year of competition and efficiency score (ratio of attained score and all possible score) ($n = 54, r = -0.71, p < 0.01$; weighted by sample size for each year) (IMO, 2016). Furthermore, the current findings are consistent with many studies that have observed decline on attentionally demand visuospatial tasks (Pietschnig & Gittler, 2015; Woodley of Menie & Fernandes, 2016). Also, it is plausible that a higher proportion of the aging population will play some part in the given negative relationship since it was found that performance on the visual-spatial tasks that recruited attentional resource decreased at faster rates than the verbal tasks as a function of age (Hale et al., 2011). Due to the fact that this is the first study that investigated the FE gain for visuospatial STM and the FE decline for visuospatial WM, further investigations are strongly suggested.

For moderating effects on mean scores of memory tests, the positive relationship between mean age and the mean scores on FDS, BDS, and FCB suggested higher performance of older than younger participants for verbal and visuospatial memory tests. However, the magnitudes of correlation, whilst significant, were small ($r_{\mathrm{FDS}} = 0.14$, $r_{\mathrm{BDS}} = 0.17$, and $r_{\mathrm{FCB}} = 0.06$, all $p < 0.01$). This finding is partially contrary to prevailing notions as to aging resulting in memory decline (León, Tascón, & Cimadevilla, 2016; Salthouse, 2009). Due to the fact that our dataset contained a high proportion of older participants and they are nearly 90% from developed countries, it is plausible that enriched environment, education, and health care systems in developed countries protected the participants in part from short-term or working memory decline due to aging processes (Fuchs et al., 2016; Leal-Galicia, Castañeda-Bueno, Quiroz-Baez, & Arias, 2008; Nouchi & Kawashima, 2014; Springer, McIntosh, Winocur, & Grady, 2005; Uchida & Kawashima, 2008). On the other hand, a negative effect of aging on memory decline was found in BCB and the magnitude of relationship is moderate ($r_{\mathrm{BCB}} = -0.45$, $p < 0.01$). This result is consistent with previous studies that found a more pronounced aging effect on visuospatial than verbal WM (Brockmole & Logie, 2013; Peich, Husain, & Bays, 2013).

Sex differences on STM and WM scores have been demonstrated in the current study. Although higher performances for female participants were observed on FDS, BDS, and FCB, the magnitudes of relationship were marginal ($r_{\mathrm{FDS}} = -0.03$, $r_{\mathrm{BDS}} = -0.01$, and $r_{\mathrm{FCB}} = -0.06$, all $ps < 0.01$). Nevertheless, this finding is supported by previous research indicating women outperformed men on verbal (Lowe et al., 2003) and visual memory measures (Andreano & Cahill, 2009). In fact, with a higher number of older adults in the current dataset, the findings are consistent with a previous study demonstrating older men doing more poorly on memory tests and having smaller hippocampal volumes than older women (DeCarli, 2015), although this might not apply to the current data in that hippocampus is thought to be mainly involved in long-term memory. In contrast, the findings for the BCB were in line with several previous studies suggesting that male participants were superior to women participants (Brunetti, Del Gatto, & Delogu, 2014b; Jones & Healy, 2006). However, this finding may also need to be interpreted with caution because of small independent samples of men and women (13 vs. 15) on BCB.

The study also found that there were higher memory scores on STM and WM tests that use computerized platforms in comparison to ‘paper and pencil’ versions. As Noyes and Garland (2008) pointed out, many paper-based or non-computerized tasks have been transferred onto computers but there may not be total equivalence. There are few comparison studies, but a better FDS score has been observed for a computer-based version in comparison to the non-computer version (Tractenberg & Freas, 2007). A converse higher numerical FCB score was also shown for the most often used wooden platform but the difference did not achieve statistical significance (Robinson & Brewer, 2016). An important point should be considered when considering the current finding is that generalization may not be warranted because of the disproportionately few independent samples using computerized platforms in our dataset, as shown in Table 1 ($\mathrm{FDS}_{\mathrm{computerized}} = 8.90\%$, $\mathrm{BDS}_{\mathrm{computerized}} = 7.10\%$, $\mathrm{FCB}_{\mathrm{computerized}} = 25.40\%$, and $\mathrm{BCB}_{\mathrm{computerized}} = 10.80\%$).

The study made a comparison between developed and developing countries, finding that people in developed countries showed higher performance on all four tests. To date, although there has been no previous meta-analysis investigating country type for working memory, a recent study by our group has indicated that participants in developed countries have significantly higher scores on general cognitive ability tests (Wongupparaj et al., 2015). Health and nutrition issues are the main concerns in many developing countries (Müller & Krawinkel, 2005) and these problems have been found to affect the memory span of children (Jukes, 2005; Miu et al., 2016; Niehaus et al., 2002). Furthermore, many studies on illiterate and unschooled people have demonstrated that low scores on STM and WM tasks were usually observed in illiterate people (Kosmidis, Zafiri, & Politimou, 2011). In all, once all moderators (age, sex, test platform and type of country) were included in the model, the relationships between year of publication and mean scores for STM and WM tasks remain statistically significant. This indicated the robustness of the FE and anti-FE on memory measures.

Finally, even with significant FE and anti-FE for verbal/visuospatial STM and for verbal/visuospatial WM in the present study, the FE and anti-FE are relatively small (*rs* = 0.12 for FDS, 0.10 for FCB, − 0.06 for BDS, and − 0.17 for BCB, respectively). Only small shifts in verbal STM and WM task performance were also found in Woodley of Menie and Fernandes (2015). Indeed, it may also be the case that the small effect may in part be attributable to the limited range of variability in test scores for the memory tasks considered. For instance, around 90% of healthy adults can recall on DS tasks somewhere between 5 and 8 digits (Wechsler, 1939). In addition, it is also plausible that other underlying reasons may account for the decline of the mean scores for BDS over time, for example, diminishing returns of IQ-boosting effects, global ceiling effects and then reversal of scores on IQ tests (Pietschnig, 2016; Pietschnig & Gittler, 2015).

5. Conclusions

In summary, verbal and visuospatial STM shows a gradual rise, supporting the FE for STM as measured by FDS and FCB, whilst verbal and visuospatial WM gradually declines over the past four decades, supporting the anti-FE for WM as indexed by BDS and BCB. Over time, environmental influences might have driven the test score changes for both types of STM but speculatively, dysgenic selection against general intelligence and also possibly age-related cognitive decline could have influenced the declining test score for verbal and visuospatial WM, especially for developed countries. These patterns of the results are in line with the predictions from the co-occurrence model, that is, the FE effect possibly occurs on less *g*-loaded abilities, whereas the anti-FE effect may concentrate on high *g*-loaded abilities.

However, these analyses only examined performances on digit span and Corsi block task to measure respectively verbal and visuospatial STM and WM. Further research should explore other standard STM and WM measures to explore and to extend the current findings. For example, this might include word span (Baddeley, 1986; Tallnd, 1965) and letter span (Kinsbourne, 1974; Taub, 1975) for verbal STM, the visual pattern test (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999) for spatial STM, reading span (Daneman & Carpenter, 1980) for verbal WM, and n-back tasks (Jaeggi, Buschkuehl, Perrig, & Meier, 2010; Kirchner, 1958) for verbal and visuospatial WM. Such studies would help to further elucidate the potential FE mechanisms of change in relation to STM and WM.

Acknowledgements

This research was supported by funds from the Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London, and the South London and Maudsley NHS Foundation Trust for some of the time of Professor Veena Kumari. Robin Morris is funded for this research by the King's College Hospital NHS Foundation Trust, London, UK.

Conflict of interest

The authors declare no conflict of interest.

Contributors

Peera Wongupparaj, Veena Kumari and Robin G. Morris designed the study. Peera and Rangsirat Wongupparaj undertook the statistical analysis, data collection and prepared the first draft. All authors contributed to and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.intell.2017.07.006.

References

Abdel-Khalek, A. M., & Lynn, R. (2008). Intelligence, family size and birth order: Some data from Kuwait. *Personality and Individual Differences*, *44*(4), 1032–1038. http://dx.doi.org/10.1016/j.paid.2007.10.029.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30–60.

Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior Genetics*, *31*(6), 615–624. http://dx.doi.org/10.1023/a:1013353613591.

Andreano, J. M., & Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learning & Memory*, *16*(4), 248–266. http://dx.doi.org/10.1101/lm.918309.

Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W. S., Swigart, A. G., & Menon, V. (2013). Visuo–spatial working memory is an important source of domain-general vulnerability in the development of arithmetic cognition. *Neuropsychologia*, *51*(11), 2305–2317. http://dx.doi.org/10.1016/j.neuropsychologia.2013.06.031.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: A longitudinal study on frequent repetitive cognitive testing. *BMC Neuroscience*, *11*(1), 1–12. http://dx.doi.org/10.1186/1471-2202-11-118.

Booth, R. W., Sharma, D., & Leader, T. I. (2016). The age of anxiety? It depends where you look: Changes in STAI trait anxiety, 1970–2010. *Social Psychiatry and Psychiatric Epidemiology*, *51*(2), 193–202. http://dx.doi.org/10.1007/s00127-015-1096-0.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(5), P223–P233. http://dx.doi.org/10.1093/geronb/60.5.P223.

Bowles, R. P., & Salthouse, T. A. (2003). Assessing the age-related effects of proactive interference on working memory tasks using the Rasch model. *Psychology and Aging*, *18*(3), 608–615.

Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75. *Frontiers in Psychology*, *4*, 12. http://dx.doi.org/10.3389/fpsyg.2013.00012.

Brunetti, R., Del Gatto, C., & Delogu, F. (2014a). eCorsi: Implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology*, *5*. http://dx.doi.org/10.3389/fpsyg.2014.00939.

Brunetti, R., Del Gatto, C., & Delogu, F. (2014b). eCorsi: Implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology*, *5*, 939. http://dx.doi.org/10.3389/fpsyg.2014.00939.

Bureau, P. R. (2016). 2015 World Population Data Sheet (Web page). Retrieved 18 June 2016 http://www.prb.org/pdf15/2015-world-population-data-sheet_eng.pdf.

Chase, W. G., & Ericsson, K. A. (1981). *Skilled memory*. Hillsdale, NJ: Erlbaum.

Chen, H.-Y., Chen, Y.-H., Liao, Y.-K., & Chen, H.-P. (2013). Relationship of fertility with intelligence and education in Taiwan: A brief report. *Journal of Biosocial Science*, *45*(04), 567–571. http://dx.doi.org/10.1017/S0021932012000545.

Cohen, J., Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. London: Sage.

Cohen, R. L., & Sandberg, T. (1977). Relation between intelligence and short-term memory. *Cognitive Psychology*, *9*(4), 534–554. http://dx.doi.org/10.1016/0010-0285(77)90019-6.

Colom, R., Abad, F. J., Quiroga, M.Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, *36*(6), 584–606.

Colom, R., Abad, F. J., Rebollo, I., & Chun Shih, P. (2005). Memory span and general intelligence: A latent-variable approach. *Intelligence*, *33*(6), 623–642.

Colom, R., Flores-Mendoza, C., Quiroga, M.Á., & Privado, J. (2005). Working memory and general intelligence: The role of short-term storage. *Personality and Individual Differences*, *39*(5), 1005–1014.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International*, *34*, 819B.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. http://dx.doi.org/10.1016/S0022-5371(80)90312-6.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103*(2), 186–201. http://dx.doi.org/10.1016/j.jecp.2009.01.004.

DeCarli, C. (2015). A call for new thoughts about what might influence human brain aging: Aging, apolipoprotein e, and amyloid. *JAMA Neurology*, *72*(5), 500–502. http://dx.doi.org/10.1001/jamaneurol.2015.33.

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo–spatial memory. *Neuropsychologia*, *37*(10), 1189–1199. http://dx.doi.org/10.1016/S0028-3932(98)00159-6.

Donnelly, K., & Twenge, J. M. (2017). Masculine and feminine traits on the Bem sex-role inventory, 1993–2012: A cross-temporal meta-analysis. *Sex Roles*, *76*(9), 556–565. http://dx.doi.org/10.1007/s11199-016-0625-y.

Dutton, E., & Lynn, R. (2013). A negative Flynn effect in Finland, 1997–2009. *Intelligence*, *41*(6), 817–820. http://dx.doi.org/10.1016/j.intell.2013.05.008.

Dutton, E., & Lynn, R. (2015). A negative Flynn effect in France, 1999 to 2008–9. *Intelligence*, *51*, 67–70. http://dx.doi.org/10.1016/j.intell.2015.05.005.

Ediev, D. M. (2011). Life expectancy in developed countries is higher than conventionally estimated. Implications from improved measurement of human longevity. *Journal of Population Ageing*, *4*(1), 5–32. http://dx.doi.org/10.1007/s12062-011-9040-x.

Flynn, J. R. (1984). The mean IQ of Americans massive gains 1932 to 1978. *Psychological*

Bulletin, *95*, 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.

Flynn, J. R. (1990). Massive IQ gains on the Scottish WISC: Evidence against brand et al.'s hypothesis. *The Irish Journal of Psychology*, *11*(1), 41–51. http://dx.doi.org/10.1080/03033910.1990.10557787.

Flynn, J. R. (2012). *Are we getting smarter: Rising IQ in the twenty-first century*. Cambridge University Press.

Fuchs, F., Cosquer, B., Penazzi, L., Mathis, C., Kelche, C., Majchrzak, M., & Barbelivien, A. (2016). Exposure to an enriched environment up to middle age allows preservation of spatial memory capabilities in old age. *Behavioural Brain Research*, *299*, 1–5. http://dx.doi.org/10.1016/j.bbr.2015.11.019.

Galton, F. (1869). *Hereditary genius*. London: MacMillan.

Ganzach, Y. (2016). Another look at the Spearman's hypothesis and relationship between Digit Span and General Mental Ability. *Learning and Individual Differences*.. http://dx.doi.org/10.1016/j.lindif.2015.12.026.

Gignac, G. E. (2015). The magical numbers 7 and 4 are resistant to the Flynn effect: No evidence for increases in forward or backward recall across 85 years of data. *Intelligence*, *48*, 85–95. http://dx.doi.org/10.1016/j.intell.2014.11.001.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & the International Child Development Steering, G (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet*, *369*(9555), 60–70. http://dx.doi.org/10.1016/S0140-6736(07)60032-4.

Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-Murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging*, *26*(1), 92–110. http://dx.doi.org/10.1037/a0021483.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.

Hester, R., Kinsella, G. J., & Ong, B. (2004). Effect of age on forward and backward span tasks. *Journal of the International Neuropsychological Society*, *10*(04), 475–481. http://dx.doi.org/10.1017/S1355617704104037.

Hill, A. C., Laird, A. R., & Robinson, J. L. (2014). Gender differences in working memory networks: A BrainMap meta-analysis. *Biological Psychology*, *102*, 18–29. http://dx.doi.org/10.1016/j.biopsycho.2014.06.008.

Howse, K. (2006). *Increasing life expectancy and the compression of morbidity: A critical review of the debate*. Oxford: Oxford Institute of Ageing, University of Oxford.

Hubber, P. J., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *Quarterly Journal of Experimental Psychology (2006)*, *67*(5), 936–954. http://dx.doi.org/10.1080/17470218.2013.838590.

IMO (2016). Result: Culmulative results by year. Retrieved from https://www.imo-official.org/results_year.aspx?column=difficulty&order=desc.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394–412. http://dx.doi.org/10.1080/09658211003702171.

Jensen, A. R., & Figueroa, R. A. (1975). Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology*, *67*(6), 882–893.

Jones, C. M., & Healy, S. D. (2006). Differences in cue use and spatial memory in men and women. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1598), 2241–2247. http://dx.doi.org/10.1098/rspb.2006.3572.

Jukes, M. (2005). The long-term impact of preschool health and nutrition on education. *Food and Nutrition Bulletin*, *26*(2 Suppl. 2), S193–S201. http://dx.doi.org/10.1177/15648265050262s210.

Kanaya, T., & Ceci, S. J. (2011). The Flynn effect in the WISC subtests among school children tested for special education services. *Journal of Psychoeducational Assessment*, *29*(2), 125–136. http://dx.doi.org/10.1177/0734282910370139.

Kanazawa, S. (2014). Intelligence and childlessness. *Social Science Research*, *48*, 157–170. http://dx.doi.org/10.1016/j.ssresearch.2014.06.003.

Karlsgodt, K. H., Bachman, P., Winkler, A. M., Bearden, C. E., & Glahn, D. C. (2011). Genetic influence on the working memory circuitry: Behavior, structure, function and extensions to illness. *Behavioural Brain Research*, *225*(2), 610–622. http://dx.doi.org/10.1016/j.bbr.2011.08.016.

Karlsgodt, K. H., Kochunov, P., Winkler, A. M., Laird, A. R., Almasy, L., Duggirala, R., … Glahn, D. C. (2010). A multimodal assessment of the genetic control over working memory. *The Journal of Neuroscience*, *30*(24), 8197–8202. http://dx.doi.org/10.1523/JNEUROSCI.0359-10.2010.

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258. http://dx.doi.org/10.1207/S15324826AN0704_8.

Khor, G. L., & Misra, S. (2012). Micronutrient interventions on cognitive performance of children aged 5-15 years in developing countries. *Asia Pacific Journal of Clinical Nutrition*, *21*(4), 471–475.

Kinsbourne, M. (1974). Cognitive deficit and the aging brain: A behavioral analysis. *The International Journal of Aging and Human Development*, *5*(1), 41–49. http://dx.doi.org/10.2190/n46c-2a38-2aqv-26tu.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358.

Kong, A., Frigge, M. L., Thorleifsson, G., Stefansson, H., Young, A. I., Zink, F., … Stefansson, K. (2017). Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences*, *114*(5), E727–E732. http://dx.doi.org/10.1073/pnas.1612113114.

Kosmidis, M. H., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence on working memory. *Archives of Clinical Neuropsychology*, *26*(7), 575–582. http://dx.doi.org/10.1093/arclin/acr063.

Kremen, W. S., Jacobsen, K. C., Xian, H., Eisen, S. A., Eaves, L. J., Tsuang, M. T., & Lyons, M. J. (2007). Genetics of verbal working memory processes: A twin study of middle-aged men. *Neuropsychology*, *21*(5), 569–580.

Kyttälä, M., & Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education*, *23*(1), 77–94.

Leal-Galicia, P., Castañeda-Bueno, M., Quiroz-Baez, R., & Arias, C. (2008). Long-term exposure to environmental enrichment since youth prevents recognition memory decline and increases synaptic plasticity markers in aging. *Neurobiology of Learning and Memory*, *90*(3), 511–518. http://dx.doi.org/10.1016/j.nlm.2008.07.005.

León, I., Tascón, L., & Cimadevilla, J. M. (2016). Age and gender-related differences in a spatial memory task in humans. *Behavioural Brain Research*, *306*, 8–12. http://dx.doi.org/10.1016/j.bbr.2016.03.008.

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Livingston, G., & Cohn, D. (2010). *Childlessness up among all women; down among women with advanced degrees*. Washington, DC: Pew Research Center.

Lowe, P. A., Mayfield, J. W., & Reynolds, C. R. (2003). Gender differences in memory test performance among children and adolescents. *Archives of Clinical Neuropsychology*, *18*(8), 865–878. http://dx.doi.org/10.1016/S0887-6177(02)00162-2.

Lynn, R. (2011). *Dysgenics: Genetic deterioration in modern populations*. London: Ulster Institute for Social Research (Second revised edition).

McNab, F., Zeidman, P., Rutledge, R. B., Smittenaar, P., Brown, H. R., Adams, R. A., & Dolan, R. J. (2015). Age-related changes in working memory and the ability to ignore distraction. *Proceedings of the National Academy of Sciences*, *112*(20), 6515–6518. http://dx.doi.org/10.1073/pnas.1504162112.

Meisenberg, G. (2010). The reproduction of intelligence. *Intelligence*, *38*(2), 220–230. http://dx.doi.org/10.1016/j.intell.2010.01.003.

Menie, M. A. W. O, Fernandes, H. B. F., José Figueredo, A., & Meisenberg, G. (2015). By their words ye shall know them: Evidence of genetic selection against general intelligence and concurrent environmental enrichment in vocabulary usage since the mid 19th century. *Frontiers in Psychology*, *6*, 361. http://dx.doi.org/10.3389/fpsyg.2015.00361.

Miu, J., Negin, J., Salinas-Rodriguez, A., Manrique-Espinoza, B., Sosa-Ortiz, A. L., Cumming, R., & Kowal, P. (2016). *Factors associated with cognitive function in older adults in Mexico*. *2016*, 9. http://dx.doi.org/10.3402/gha.v9.30747.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology. General*, *130*(4), 621–640.

Müller, O., & Krawinkel, M. (2005). Malnutrition and health in developing countries. *Canadian Medical Association Journal*, *173*(3), 279–286. http://dx.doi.org/10.1503/cmaj.050342.

Myerson, J., Emery, L., White, D. A., & Hale, S. (2003). Effects of age, domain, and processing demands on memory span: Evidence for differential decline. *Aging, Neuropsychology, and Cognition*, *10*(1), 20–27. http://dx.doi.org/10.1076/anec.10.1.20.13454.

Nargund, G. (2006). Declining birth rate in developed countries: A radical policy re-think is required. *Facts Views Vision Obgyn*, *1*(3), 191–193.

Niehaus, M. D., Moore, S. R., Patrick, P. D., Derr, L. L., Lorntz, B., Lima, A. A., & Guerrant, R. L. (2002). Early childhood diarrhea is associated with diminished cognitive function 4 to 7 years later in children in a northeast Brazilian shantytown. *The American Journal of Tropical Medicine and Hygiene*, *66*(5), 590–593.

Nouchi, R., & Kawashima, R. (2014). Improving cognitive function from children to old age: A systematic review of recent smart ageing intervention studies. *Advances in Neuroscience*, *2014*, 15. http://dx.doi.org/10.1155/2014/235479.

Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, *51*(9), 1352–1375. http://dx.doi.org/10.1080/00140130802170387.

Peach, H., Lyerly, J. E., & Reeve, C. L. (2014). Replication of the Jensen effect on dysgenic fertility: An analysis using a large sample of American youth. *Personality and Individual Differences*, *71*, 56–59. http://dx.doi.org/10.1016/j.paid.2014.07.017.

Peich, M.-C., Husain, M., & Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory. *Psychology and Aging*, *28*(3), 729–743. http://dx.doi.org/10.1037/a0033236.

Pelosi, L., & Blumhardt, L. D. (1999). Effects of age on working memory: An event-related potential study. *Cognitive Brain Research*, *7*(3), 321–334. http://dx.doi.org/10.1016/S0926-6410(98)00035-4.

Pietschnig, J. (2016). The Flynn effect: Technology may be part of it, but is most certainly not all of it. *Measurement: Interdisciplinary Research and Perspectives*, *14*(2), 70–73. http://dx.doi.org/10.1080/15366367.2016.1171612.

Pietschnig, J., & Gittler, G. (2015). A reversal of the Flynn effect for spatial perception in German-speaking countries: Evidence from a cross-temporal IRT-based meta-analysis (1977–2014). *Intelligence*, *53*, 145–153. http://dx.doi.org/10.1016/j.intell.2015.10.004.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science*, *10*(3), 282–306. http://dx.doi.org/10.1177/1745691615577701.

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Pervasiveness of the IQ rise: A cross-temporal meta-analysis. *PloS One*, *5*(12), e14406. http://dx.doi.org/10.1371/journal.pone.0014406.

Robinson, S. J., & Brewer, G. (2016). Performance on the traditional and the touch screen, tablet versions of the Corsi block and the tower of Hanoi tasks. *Computers in Human Behavior*, *60*, 29–34. http://dx.doi.org/10.1016/j.chb.2016.02.047.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. http://dx.doi.org/10.1111/j.1467-9280.2006.01693.x.

Rönnlund, M., Carlstedt, B., Blomstedt, Y., Nilsson, L.-G., & Weinehall, L. (2013). Secular trends in cognitive test performance: Swedish conscript data 1970–1993. *Intelligence*, *41*(1), 19–24. http://dx.doi.org/10.1016/j.intell.2012.10.001.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis publication bias in meta-analysis*. John Wiley & Sons, Ltd1–7.

Runquist, E. A. (1936). Intelligence test scores and school marks in 1928 and 1933. *School & Society*, *43*, 301–304.

Salthouse, T. A. (1994). The aging and working memory. *Neuropsychology*, *8*(4), 535–543.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507–514. http://dx.doi.org/10.1016/j.neurobiolaging.2008.09.023.

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.

Springer, M. V., McIntosh, A. R., Winocur, G., & Grady, C. L. (2005). The relation between brain activity during memory tasks and years of education in young and older adults. *Neuropsychology*, *19*(2), 181–192. http://dx.doi.org/10.1037/0894-4105.19.2.181.

St Clair-Thompson, H. L., & Holmes, J. L. (2008). *Improving short-term and working memory: Methods of memory training*. New York: Nova Science.

Sundet, J. M., Borren, I., & Tambs, K. (2008). The Flynn effect is partly caused by changing fertility patterns. *Intelligence*, *36*(3), 183–191. http://dx.doi.org/10.1016/j.intell.2007.04.002.

Tallnd, G. A. (1965). Three estimates of the word span and their stability over the adult years. *Quarterly Journal of Experimental Psychology*, *17*(4), 301–307. http://dx.doi.org/10.1080/17470216508416448.

Taub, H. A. (1975). Mode of presentation, age, and short-term memory. *Journal of Gerontology*, *30*(1), 56–59. http://dx.doi.org/10.1093/geronj/30.1.56.

te Nijenhuis, J., & van der Flier, H. (2013). Is the Flynn effect on g? A meta-analysis. *Intelligence*, *41*(6), 802–807. http://dx.doi.org/10.1016/j.intell.2013.03.001.

Tractenberg, R. E., & Freas, C. E. (2007). Benchmarking a computerized test of immediate verbal memory. *Behavior Research Methods*, *39*(4), 863–869. http://dx.doi.org/10.3758/bf03192980.

Trahan, L., Stuebing, K. K., Hiscock, M. K., & Fletcher, J. M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin*, *140*(5), 1332–1360. http://dx.doi.org/10.1037/a0037173.

Twenge, J. M. (2000). The age of anxiety? The birth cohort change in anxiety and neuroticism, 1952–1993. *Journal of Personality and Social Psychology*, *79*(6), 1007–1021. http://dx.doi.org/10.1037/0022-3514.79.6.1007.

Twenge, J. M., & Campbell, S. M. (2008). Generational differences in psychological traits and their impact on the workplace. *Journal of Managerial Psychology*, *23*(8), 862–877. http://dx.doi.org/10.1108/02683940810904367.

Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review*, *5*(4), 321–344. http://dx.doi.org/10.1207/s15327957pspr0504_3.

Uchida, S., & Kawashima, R. (2008). Reading and solving arithmetic problems improves cognitive functions of normal aged people: A randomized controlled study. *Age*, *30*(1), 21–29. http://dx.doi.org/10.1007/s11357-007-9044-x.

van der Ven, S. H. G., van der Maas, H. L. J., Straatemeier, M., & Jansen, B. R. J. (2013). Visuospatial working memory and mathematical ability at different ages throughout primary school. *Learning and Individual Differences*, *27*, 182–192. http://dx.doi.org/10.1016/j.lindif.2013.09.003.

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. http://dx.doi.org/10.1002/jrsm.11.

Wechsler, D. (1939). *The measurement and appraisal of adult intelligence*. Baltimore, MD: The Williams & Wilkins Company.

Wechsler, D. (1981). *The Wechsler adult intelligence scale-revised manual*. San Antonio: Psychological Corporation.

Wechsler, D. (1987). *The Wechsler memory scale-revised*. San Antonio: Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale* (4th ed.). San Antonio: Psychological Corporation.

Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., van Baal, G. C. M., Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, *32*(5), 509–537. http://dx.doi.org/10.1016/j.intell.2004.07.002.

Wingfield, A., Stine, E. A. L., Lahar, C. J., & Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research*, *14*(2), 103–107. http://dx.doi.org/10.1080/03610738808259731.

Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). A cross-temporal meta-analysis of Raven's progressive matrices: Age groups and developing versus developed countries. *Intelligence*, *49*, 1–9. http://dx.doi.org/10.1016/j.intell.2014.11.008.

Woodley of Menie, M. A (2015). How fragile is our intellect? Estimating losses in general intelligence due to both selection and mutation accumulation. *Personality and Individual Differences*, *75*, 80–84. http://dx.doi.org/10.1016/j.paid.2014.10.047.

Woodley of Menie, M. A, & Dunkel, C. S. (2015). In France, are secular IQ losses biologically caused? A comment on Dutton and Lynn (2015). *Intelligence*, *53*, 81–85. http://dx.doi.org/10.1016/j.intell.2015.08.009.

Woodley of Menie, M. A, & Fernandes, H. B. F. (2015). Do opposing secular trends on backwards and forwards digit span evidence the co-occurrence model? A comment on Gignac (2015). *Intelligence*, *50*, 125–130. http://dx.doi.org/10.1016/j.intell.2015.03.003.

Woodley of Menie, M. A, & Fernandes, H. B. F. (2016). Showing their true colours: Possible secular declines and a Jensen effect on colour acuity — More evidence for the weaker variant of Spearman's Other Hypothesis. *Personality and Individual Differences*, *88*, 280–284. http://dx.doi.org/10.1016/j.paid.2015.09.009.

Woodley of Menie, M. A, & Meisenberg, G. (2013). A Jensen effect on dysgenic fertility: An analysis involving the National Longitudinal Survey of Youth. *Personality and Individual Differences*, *55*(3), 279–282. http://dx.doi.org/10.1016/j.paid.2012.05.024.

Woodley of Menie, M. A, Fernandes, H. B. F., Figueredo, A. J., & Meisenberg, G. (2015). By their words ye shall know them: Evidence of genetic selection against general intelligence and concurrent environmental enrichment in vocabulary usage since the mid 19th century. *Frontiers in Psychology*, *6*. http://dx.doi.org/10.3389/fpsyg.2015.00361.

Woodley of Menie, M. A, Figueredo, A. J., Dunkel, C. S., & Madison, G. (2015). Estimating the strength of genetic selection against heritable g in a sample of 3520 Americans, sourced from MIDUS II. *Personality and Individual Differences*, *86*, 266–270. http://dx.doi.org/10.1016/j.paid.2015.05.032.

Woodley of Menie, M. A, Peñaherrera, M. A., Fernandes, H. B. F., Becker, D., & Flynn, J. R. (2016). It's getting bigger all the time: Estimating the Flynn effect from secular brain mass increases in Britain and Germany. *Learning and Individual Differences*, *45*, 95–100. http://dx.doi.org/10.1016/j.lindif.2015.11.004.

Woodley, M. A., & Meisenberg, G. (2013). In the Netherlands the anti-Flynn effect is a Jensen effect. *Personality and Individual Differences*, *54*(8), 871–876. http://dx.doi.org/10.1016/j.paid.2012.12.022.

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., … Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology*, *33*(1), 101–111. http://dx.doi.org/10.1080/13803395.2010.493149.

Zhang, T., Tan, H., Wu, Y., Han, B., & Wang, T. (2016). Urban older adults becoming unhealthier in modern China: A cross-temporal meta-analysis. *Psychological Reports*, *118*(3), 737–747. http://dx.doi.org/10.1177/0033294116646160.